A genomic control odyssey

Isabelle S. Peter

www.elsevier.com/locate/developmentalbiology

PII: S0012-1606(16)00022-1
DOI: http://dx.doi.org/10.1016/j.ydbio.2016.01.027
Reference: YDBIO6997

To appear in: *Developmental Biology*

Cite this article as: Isabelle S. Peter, A genomic control odyssey, *Developmental Biology,* http://dx.doi.org/10.1016/j.ydbio.2016.01.027

A genomic control odyssey

Isabelle S. Peter

Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA

Tel: +1 626 395 5887. E-mail: ipeter@caltech.edu

While writing this article, I came to realize that it is utterly impossible to capture the spirit, scientific contributions, or colorful personality of Eric Davidson in a few pages, or even a whole journal, and so I won't even try. I would nevertheless like to pay tribute to an enormously powerful mind and a great intellectual partner and friend, with whom I worked very closely over the last almost nine years, and I will do so in two ways. On the one hand, I will give a brief and entirely subjective account of what I would regard as the most profound expression of his scientific intellect, which is his lifelong journey to understand the genomic control of development. On the other hand, I will remember Eric's very expressive personality in the form of but a few examples of our interactions. What impressed me most about Eric's spirit, of which I am reminded by the photograph in Figure 1, is that despite the incredible amount of his achievements, Eric never stopped until his very last day to think about what lays ahead, driven by immense curiosity and a deep desire to understand the natural world around us.

When I first met Eric in 2006 for an interview, I would not say that we got along right away. He was extremely knowledgeable, extremely smart, extremely acute, and with no patience whatsoever for formalities. Thus the polite interchange one usually associates with such occasions did not last very long, an early indication of what would follow. After listening for about 30 seconds to my past experiences, he said "oh, how boring!". As intended, this broke the ice immediately, and after a short moment of furious silence from my side, he added with a broad smile "well, we basically do very similar things here". This was of course a complete understatement, but it allowed us to continue the conversation in a productive and definitely more pleasant manner. A few months later I arrived in Pasadena and started to work in Eric's lab. In our first conversation, Eric asked me about what type of project I would be interested in, to which I enthusiastically answered that I wanted to build a computational model for the gene regulatory network they had recently analyzed. To this Eric immediately replied: "first you learn how to walk, then you can run!" And so I did, and it took about two years and one major argument until we started to engage in a scientific discourse that would last for many years and that would profoundly affect my way of thinking.

As Eric mentioned on several occasions, one of the major influences on his vision was an early encounter with Chapter XIV of E.B. Wilson's "The Cell in Development and Heredity" (Wilson, 1924), which he had to read as a thesis assignment. Himself a great visionary, Wilson stated at the beginning of the 20th century: "*How are the operations of development so coordinated as to give rise to a definitely ordered system? It is our scientific habit of thought to regard the operation of any specific system as determined primarily by its specific physical-chemical composition…. This mechanistic assumption implies some specific structure of material configuration in the system, and since the organization of the egg is hereditary, the structure or configuration must be preserved by cell division without loss of its specific character…* " . These sentences are of very much the same flavor as the questions that dominated Eric's lifelong work and with which he often introduced his own talks. To translate E.B. Wilson's words into a language Eric might have used concerning the nature of the genomic control system: Since it is clear that the control of development is encoded in the genome, and thus preserved through cell division, what is the particular configuration of this code? What are the molecular components of this system, which will read and execute the genomically encoded program? What are the mechanisms controlling the increasing complexity of spatial organization during the highly reproducible development of any given body plan?

Several hallmarks reflect Eric's journey to address these questions. His visionary work with Roy Britten, at an early stage of this long lasting intellectual partnership, produced a theoretical model which

captured many of the features which later were indeed found in gene regulatory networks, even though the model was produced at a time when the molecular players were not yet known (Britten and Davidson, 1969). The main import of this model, which is that the sequence-specific regulation of gene expression has to be key to the genomic control of developmental process, provided a first plausible solution to all three questions. Nevertheless it took several decades until it finally became possible to experimentally address genomic regulatory networks. In the meantime, Eric followed an argument that according to his own account profoundly shaped his conception of gene regulatory networks. In the early 90's, he undertook an attempt to define common regulatory principles for animal development, by comparing developmental processes across several bilaterian species (Davidson, 1990). From this analysis he deduced the type of information that had to be encoded within *cis*-regulatory sequence, as shown in Figure 2. In the following, Eric's lab tested these ideas by experimentally dissecting several *cis*-regulatory modules that contain sufficient information to drive precise spatial gene expression during development. Among these, the regulatory sequences driving expression of the cytoskeletal actin gene *cyiiia*, and of the midgut differentiation gene *endo16*, have been experimentally explored at a level of detail that even nowadays is only rarely achieved (Hough-Evans et al., 1990; Kirchhamer et al., 1996; Yuh et al., 1998; Yuh and Davidson, 1996). Of course, the results showed that actual *cis*-regulatory information is not quite as neatly organized as Eric first envisioned, but that remains a detail in comparison to the enormous insights that were acquired while testing this first model of developmental regulation of gene transcription.

The first model for the gene regulatory network specifying endomesodermal cell fates in the sea urchin embryo was published in 2002, and this opened the door for an entirely new way of thinking about the genomic control of development (Davidson et al., 2002). This time, the key molecular players were known, and so were the genomic sequences recognized by them. The main argument for constructing a model for gene regulatory networks was that if the spatial expression of differentiation genes is encoded in their *cis*-regulatory sequences, and read by specific transcription factors, then the same mechanism had to be responsible for expression of these transcription factors in the right domain of the embryo at the right time in development. Thus the expression of all regulatory genes encoding transcription factors and signaling molecules must be controlled by mutual transcriptional regulation. The network constellation of these interactions is then a result of the combinatorial regulation of *cis*-regulatory modules by multiple transcription factors, and of the multiple target genes regulated by each transcription factor. It turned out that even though the predicted regulatory functions such as "amplitude controller" or "lineage identifier" or "temporal regulator" are all indeed impacting the activity of *cis*-regulatory modules, these functions are much more intertwined and executed by some of the same transcription factors. The precise role of any transcription factor providing input into a given *cis*-regulatory module is now no longer assumed to be a function of a particular property of that molecule, but a function of the position of this regulatory gene within the gene regulatory network circuitry.

Eric, together with a large group of collaborators without whom this enormous achievement would have never seen the light of day, had thus arrived at a solution which explained the progressive developmental specification of different fates as a function of regulatory genomic sequences. And this point, both the genomic sequences encoding this developmental program, as well as the molecules reading and executing it, were experimentally accessible, as was demonstrated by the first gene regulatory network model. To the most part, these network models were designed to explain the

differential specification of spatial domains based on identical genomic sequence. But at the same time, the processes of transcription factor binding to particular cis-regulatory sequences, transcription, and translation, are all processes that operate in real time. Together with Hamid Bolouri, Eric modeled the dynamics of a gene regulatory cascade, which calculated based on known parameters, the time it would take from the initial transcription of an upstream regulatory gene, to the activation of its target gene (Bolouri and Davidson, 2003). The time delay calculated in this model was very much consistent with measured gene expression data, which demonstrated that gene regulatory networks could not only explain the spatial compartmentalization of an embryo, but also the temporal progression of developmental processes.

My first project in the lab was to elaborate the network underlying endoderm specification. Since by that time, the sea urchin genome was sequenced, and the expression of regulatory genes in individual embryonic domains was analyzed, it became possible to revise the initial 2002 model and to obtain a more complete characterization of this network. Once I had generated sufficient data to start building a model for the endodermal network, I came up with what I thought to be a great solution for how this network specifies the early endodermal domains. The only problem was that it directly contradicted the models Eric had in mind at that time. Eric and I managed to address the discrepancy in one short, heated argument which was so disruptive that we did not exchange another word for the next couple of months. Eventually this situation became so boring and silly that we decided to move on to a more productive mode of interaction. We agreed to meet again to discuss the subject. This time, instead of confronting our views, we focused on collecting every piece of evidence and every argument we could find. One by one, we worked through a pile of data, and asked what type of experiment the data came from, what they might mean and what not, and how confident we could be in that conclusion. At this point we started to synthesize a new model, one we both could accept as the best solution, because it was built based on all available evidence and the most comprehensive analysis we could perform at the time.

Curiously, Eric and I both considered this initial dispute as formative for our collaboration, and we often laughed about it once sufficient time had passed. We would repeat a similar procedure for every new project, although our arguments became much friendlier over time. Once we thought of a new and exciting problem to work on, we would each think about possible ways to solve it. Eric would of course immediately write down a properly structured outline of the problem, including the individual steps it would take to solve it, while I was far less organized and just came to the meeting with some ideas in my mind. What we had in common though is that we were both convinced that the solution we had each developed would be the only logical way to address the problem. It never failed to surprise that we had come up with completely different solutions. Sometimes these solutions were so different that we had to question whether we were even trying to address the same problem. Although over time we learned to keep our arguments within a tolerable range, I vividly remember the pleasantries that were exchanged in the initial phase of every project. One of Eric's favorite answers to my complaints about logical inconsistencies was "consistency is the hobgoblin of small minds", spoken with such deep conviction that I immediately felt like a small minded hobgoblin. I would eventually learn to initiate my arguments with "I'm sorry for being so small minded, but …", just to avoid being transformed into a silly creature. If on the other hand Eric would agree with something I said, this was frequently commented with "I think that is correct, that's why I wrote it in the 70's…". I loved Eric's sense of humor. To be fair, my own statements were not exclusively appreciative either, but I will restrain from reproducing them

here. What was important however, was that by the end of this exercise we had deflated both arguments and egos, and were ready to construct a new solution which would incorporate elements of what each of us had brought into the discussion. The excitement over the new solution, which we both agreed was much better then what either of us had in mind initially, was entirely overwhelming. By then, we had completely forgotten about the painful process it took to get there, and we would jump right into the next project.

One of the major projects worked out like this was the project to build a Boolean computational model of the endomesoderm network. This was in 2011 when the gene regulatory networks for several domains of the sea urchin embryo had been experimentally solved to incorporate all known regulatory genes. But how would one know whether the solution of the network was complete? How could it be measured? We first started to identify the principle features determining the operation of gene regulatory networks, and came to the conclusion that the single most important assumption to be tested is that the total *cis*-regulatory information associated with all regulatory genes in the system would be sufficient to define the spatial and temporal patterns of gene expression driving developmental progression. Since in addition, we had a long list of gene regulatory network features that would have to be captured by a computational model that were not part of any available computational approach, we started to design our own Boolean model. As usual, the exercise started with a piece of paper and a pen, and we built a pilot Boolean model for a small subcircuit, just to test whether we could get this to work. Once we did, we went systematically through the entire network, again revisiting the underlying experimental evidence, until every part of the network was captured in countless Boolean logic statements. The actual computational model was then programmed by Emmanuel Faure, who very cheerfully dealt with the endless requests to modify or at one point even completely revise the model. It was almost unbelievable when the Boolean model was finalized and we realized that the endomesoderm network model indeed contained a nearly complete set of instructions for developmental gene expression for the first 30h of sea urchin embryogenesis (Peter et al., 2012). This Boolean model provided a proof that the gene regulatory networks encoded in genomic regulatory sequences capture developmental progression. For Eric, this meant that the concepts he had pioneered indeed provided the molecular solution to the set questions he had started out with. For me, I was lucky to have been at the right place at the right time to be part of this adventure which has so profoundly shaped my view of the genomic code for development.

The last and most intense part of our collaboration was writing our book “Genomic Control Process, Development and Evolution”, Eric’s sixth book and my first one (Peter and Davidson, 2015). Inspired by the success of the Boolean model, Eric decided it would be time to write another book, but only if I would agree to be a co-author. I did not think twice, since I considered this as a once in a lifetime opportunity, although I sometimes came to wish I had given it a little more thought. It was beyond my imagination how much work it would be to write a book of this kind, in which large areas of biology are incorporated into one logically consistent (!) framework. Again, the approach was entirely systematic, as was typical for Eric, and we first produced an outline for the entire book. Our idea at first was to split up the entire book into separate sections, discuss what will be covered in each one, and then work separately on drafting these sections. However, over time we became very experienced in writing together, which saved us substantial time and became an extremely productive and stimulating way of collaborating. It never ceased to surprise me how much Eric knew about an enormous array of topics, and how he nevertheless remained open to think about them in entirely new ways. Despite being over

75, Eric had no problem working long days with an intensity and energy that was demanding to keep up with at any age. Nevertheless, it took us almost two years to write this book, which was well beyond the six months Eric had initially projected. I was never quite sure whether Eric was serious when he said “in the unlikely case that we ever were to finish this book…”, but as time went on, it became clear that it had to be finished sooner rather than later. As we wrote the very last sentences of the book, it made me realize that for Eric it represented a final conclusion of his long quest. To arrive at a satisfactory level of understanding by following the same set of questions for such a long and extremely productive scientific life must be one of the most rewarding experiences I can think of. It is this entire journey that very much reflects the focus and strength of Eric’s marvelous scientific mind. Needless to say, the science Eric shaped so profoundly will be continued by the many scientists whose life he touched by providing endless inspiration. As a person though, Eric will be dearly missed.

References

Bolouri, H., Davidson, E.H., 2003. Transcriptional regulatory cascades in development: initial rates, not steady state, determine network kinetics. Proc. Natl Acad. Sci. USA 100, 9371-9376.
Britten, R.J., Davidson, E.H., 1969. Gene regulation for higher cells: a theory. Science (New York, N.Y 165, 349-357.
Davidson, E.H., 1990. How embryos work: a comparative view of diverse modes of cell fate specification. Development (Cambridge, England) 108, 365-389.
Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., Otim, O., Brown, C.T., Livi, C.B., Lee, P.Y., Revilla, R., Rust, A.G., Pan, Z., Schilstra, M.J., Clarke, P.J., Arnone, M.I., Rowen, L., Cameron, R.A., McClay, D.R., Hood, L., Bolouri, H., 2002. A genomic regulatory network for development. Science (New York, N.Y 295, 1669-1678.
Hough-Evans, B.R., Franks, R.R., Zeller, R.W., Britten, R.J., Davidson, E.H., 1990. Negative spatial regulation of the lineage specific CyIIIa actin gene in the sea urchin embryo. Development (Cambridge, England) 110, 41-50.
Kirchhamer, C.V., Yuh, C.H., Davidson, E.H., 1996. Modular cis-regulatory organization of developmentally expressed genes: two genes transcribed territorially in the sea urchin embryo, and additional examples. Proc. Natl Acad. Sci. USA 93, 9322-9328.
Peter, I.S., Davidson, E.H., 2015. Genomic Control Process, Development and Evolution. Academic Press/Elsevier.
Peter, I.S., Faure, E., Davidson, E.H., 2012. Feature Article: Predictive computation of genomic logic processing functions in embryonic development. Proc Natl Acad Sci U S A 109, 16434-16442.
Wilson, E.B., 1924. The Cell in Development and Heredity, 3rd ed. Macmillan, New York.
Yuh, C.H., Bolouri, H., Davidson, E.H., 1998. Genomic cis-regulatory logic: experimental and computational analysis of a sea urchin gene. Science (New York, N.Y 279, 1896-1902.
Yuh, C.H., Davidson, E.H., 1996. Modular cis-regulatory organization of Endo16, a gut-specific gene of the sea urchin embryo. Development (Cambridge, England) 122, 1069-1082.

Figure Captions

Figure 1: Eric Davidson always on the lookout for something new, or as here, enjoying a game of his favorite football team, the San Diego Chargers.

Figure 2: An early definition of developmental regulatory functions that must be encoded in cis-regulatory sequence (from Davidson, 1990). The figure is best described by an excerpt of the original caption: *"Aspects of regulatory architectdure relevant to early development. (A) The 'smart' histospecific structural gene [i.e. differentiation gene]: A nonhierarchical model for developmental control of gene expression. The regulatory system behaves like a logic chip and integrates the state of all of the different inputs. Each input (i.e. species of active factor interaction) itself organizes a set of genes that share its binding site. Thus there will be sets of genes that respond to C, the cell cycle operator; a different set that responds to T, the temporal operator;…"* . (B) The sea urchin embryo: scheme of an early sea urchin embryo (left) showing the territories known at the time to be specified during early embryogenesis, and the parts of the sea urchin larva (right) they give rise to. Open, oral ectoderm; hatched, aboral ectoderm; wavy line, gut and definitive mesenchyme; solid, skeletogenic mesenchyme.

WELLS FARGO

Figure 2